

# Algoritmo Cromático Aplicado A Un Modelo De Regresión No Lineal En Pronósticos De Series De Tiempo

*Chromatic Algorithm Applied To A Non-Linear Regression Model In Time Series Forecasts*

Rafael Sabie Lobo.<sup>1\*</sup>   
Universidad de Córdoba

Jose Aviles Pacheco<sup>2</sup>   
Universidad de Córdoba

Jorge Gómez Gómez<sup>3</sup>   
Departamento de ingeniería de Sistemas, Universidad de Córdoba

Fecha recepción: 15 de junio de 2020

Fecha aceptación: 16 de julio de 2020

© 2020 Universidad de Córdoba. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Attribution License, que permite el uso ilimitado, distribución y reproducción en cualquier medio, siempre que el autor original y la fuente se acreditan.

## RESUMEN

El presente trabajo utiliza la nueva metaheurística, algoritmo cromático para la optimización de pronósticos de series de tiempo a través de un modelo de regresión no lineal. En esta nueva propuesta se maneja el algoritmo cromático debido a sus características de codificación real y a su memoria de arranque múltiple, que le permiten ser más eficiente a la hora minimizar el error en los pronósticos del

## ABSTRACT

The present work uses the new metaheuristic, a chromatic algorithm for the optimization of time series forecasts through a non-linear regression model. In this new proposal, the chromatic algorithm is handled due to its real coding characteristics and its multiple boot memory, which allow it to be more efficient when minimizing the error in the model's forecasts, to achieve this

<sup>1</sup> Estudiantes de Ingeniería de sistemas

<sup>2</sup> Estudiantes de ingeniería de sistemas

<sup>3</sup> Doctorado en Tecnologías de la Información y la Comunicación, docente Universidad de Córdoba

modelo, para lograr esto se utilizan indicadores estadísticos del error que contribuyen a mejorar las predicciones a cada problema específico. Además, se idean mejoras al modelo de regresión y al algoritmo utilizado de tal manera que se logra predecir el comportamiento de los problemas, no solo de una variable sino también de múltiples variables. El algoritmo junto con el modelo es probado en distintos problemas de una y múltiples variables proporcionando muy buenas predicciones.

Se ejecuta en caso de estudio práctico relacionado con la estimación de los precios de ganado según su tipo en la región estudiada. Este nuevo método genera más posibilidades para alcanzar que los pronósticos se ajusten y para mejorar cualquier tipo de predicción. Esta investigación proporciona una nueva manera para minimizar los errores en los pronósticos y generar resultados de gran calidad.

También demuestra que se permite establecer pronósticos tanto en problemas de una variable como en los de múltiples variables, con tiempos computacionales razonables. Esta sería una excelente estrategia para las innumerables empresas, entidades u organizaciones que requieren métodos verdaderamente eficientes que les permitan tomar las mejores decisiones.

**PALABRAS CLAVE:** Pronósticos, metaheurísticas, series de tiempo.

statistical indicators of the error are used that help improve predictions for each specific problem.

In addition, improvements are devised to the regression model and the algorithm used in such a way that it is possible to predict the behavior of the problems, not only of one variable but also of multiple variables. The algorithm together with the model is tested in different problems with one and multiple variables, providing very good predictions. It is executed in a practical case study related to the estimation of livestock prices according to their type in the studied region. This new method generates more possibilities to achieve that the forecasts are adjusted and to improve any type of prediction. This research provides a new way to minimize forecast errors and generate high-quality results.

It also shows that it is possible to establish forecasts in both single and multi-variable problems, with reasonable computational times. This would be an excellent strategy for the countless companies, entities or organizations that require truly efficient methods that allow them to make the best decisions.

**KEYWORDS:** Forecasts, metaheuristics, time series.

## INTRODUCCIÓN

En el transcurrir del tiempo muchos tipos de metaheurísticas surgido con el fin de ser usadas para mejorar diferentes tipos de problemas de gran complejidad o magnitud. Estos métodos de van desde búsquedas de vecindad hasta las evolutivas, así como la combinación de las mismas.

A continuación, se mencionan algunas de los métodos de búsqueda más importantes y representativos de los últimos años.

Con relación al análisis de series de tiempo para la predicción de los precios del ganado Córdoba - Colombia se identifica en el año 2007 el trabajo “Estacionalidad, ciclos y volatilidad en los precios del ganado macho de levante en Montería, Colombia” realizado por Omar Castillo (Castillo, Estacionalidad, ciclos y volatilidad en los precios del ganado macho de levante en Montería, Colombia, 2007). En su desarrollo se describe el comportamiento temporal de los precios del ganado vivo macho de levante de primera calidad en la ciudad de Montería, Colombia comercializado en las subastas. Materiales y métodos. Se realizaron análisis de los precios semanales y mensuales durante el período 1997-2006 utilizando técnicas estadísticas y econométricas como la media móvil multiplicativa, la tasa de crecimiento sobre medias anuales, 12 T12, y modelos auto-regresivos heterocedásticos condicionales, ARCH, o GARCH. Como resultado se denota evidencias de estacionalidad y ciclos en los precios mensuales; no hubo evidencia de comportamientos volátiles en precios semanales de los ganados de 1,1¼ y 1½ años de edad, pero si para los de 1 año.

En este mismo año también se encontró el trabajo “Comportamiento temporal de los precios del ganado macho de levante de primera en Sincelejo” (Castillo, Comportamiento temporal de los precios del ganado macho de levante de primera en Sincelejo, 2007). De manera similar el autor

Omar Castillo señala que igualmente el comportamiento temporal de los precios del ganado vivo macho de levante de primera calidad en la ciudad de Sincelejo, comercializado en las subastas. Para ello se acude al análisis de los precios mensuales y semanales durante el período 1997 - 2006 utilizando técnicas estadísticas y econométricas como la media móvil multiplicativa, la tasa de crecimiento sobre medias anuales, 12 T12, y modelos auto-regresivos heterocedásticos condicionales, ARCH, o GARCH. Los resultados indican la presencia de estacionalidad y ciclos en los precios mensuales; en los precios semanales de las edades de 1 y de 1¼ año se observaron comportamientos volátiles, el precio de los de 1½ año se han movido dentro de rangos estables.

Seguidamente se destaca el trabajo realizado por Botero y Cano en 2008, en su trabajo “Análisis de series de tiempo para la predicción de los precios de la energía en la bolsa de Colombia”. En este se realiza un análisis del comportamiento del precio de la energía eléctrica ha incrementado su volatilidad, reflejando el riesgo existente para los diferentes agentes que intervienen en el mercado. El objetivo de este artículo es presentar una metodología para la implementación de modelos de regresión, sobre la serie histórica de precios de bolsa de energía en Colombia (Botero & Cano, 2008).

En el año 2014 se identifica la investigación “Uso de un modelo univariado de series de tiempo para la predicción, en el corto plazo, del comportamiento de la producción de carne de bovino en Baja California, México”. Esta fue desarrollada por Barreras, Sánchez y Figueroa en Baja California, México. Se utilizó el modelo autorregresivo de promedios móviles ARMA. Los resultados respaldan el modelo para obtener a corto plazo predicciones de producción de carne en baja california (Barreras Serrano, Eduardo, & Figueroa Saavedra, 2014).

De manera más reciente de encuentra la investigación titulada “Descripción y pronóstico de la producción de carne de bovino en el estado de Tabasco”. Esta misma fue efectuada en el año 2017 por Espinosa, Vélez, Quiroz, Granados, Moctezuma y Casanova en México obedeciendo los siguientes objetivos. Analizar el comportamiento y productividad de la producción de carne de bovino en Tabasco y encontrar un modelo de series de tiempo capaz de realizar un pronóstico de la producción, confiable, cercano a los valores de la serie, cuya correlación sea estadísticamente significativa, se integró una serie de tiempo anual de producción de carne en canal (PCaCa), animales sacrificados (ASac) y peso en canal (PeCa) de los ASac en los 17 municipios del estado para el periodo 2006-2015. También se integró una serie mensual de PCaCa de enero de 2003 a marzo de 2017. Bajo este precedente se encontró que Tabasco produce en promedio por año 65 338 ton de carne con 310 564 ASac con un peso promedio de 211 kg.

También se encontró que tanto la PCaCa, como los ASac presentaron TMAC positivas para el estado, siendo el DDR de Villahermosa el de mayor dinamismo, con una tasa de crecimiento promedio de 2 %; y el municipio de Teapa el de a mayor TCMA, con 4,4 %. El modelo de serie de tiempo que mejor predice la PCaCa fue el autorregresivo integrado de promedio móvil, ARIMA (p, d, q) (P, D, Q) k, de orden 1, por tanto, la producción de carne en Tabasco tiene un comportamiento estacional. Se concluye que el modelo seleccionado permite hacer un pronóstico cercano a la realidad con una confiabilidad estadística significativa. (Espinosa, Vélez, & A., 2017)

Revisando lo más reciente en la literatura del ámbito colombiano encontramos el trabajo “Determinantes del precio del ganado

gordo bovino en pie de Medellín y Bogotá: 2009-2019”. Este fue elaborado por Sebastián López Nieto en la Universidad EAFIT - Escuela De Economía y Finanzas. Este trabajo deja en evidencia las dinámicas, entre 2009 y 2019, que rigen los precios del ganado bovino macho y hembra entre los principales mercados de carne en Colombia; ya que permite entender de mejor manera el mercado bovino como oferta final de los frigoríficos, lo cual afecta directamente las decisiones de gran cantidad de agentes que comercian en él. Entender la dinámica de este mercado, termina impactando positivamente los ganaderos y los frigoríficos. (Lopez, 2019)

Los ganaderos al tener más y mejor información, pueden mejorar su productividad, beneficiando su calidad de vida. Por parte de los frigoríficos, con mejor información se puede impulsar de mejor manera la competitividad del sector, que promete ser clave en los próximos años en materia de exportaciones. Además, con lo hablado con ambos frigoríficos, es de gran información estos estudios acerca de los precios y es una manera como estos puedan ayudarles a tomar mejores decisiones a los agentes.

Des la óptica de lo investigado se señala que es claro que para determinar el precio del ganado bovino macho hay 3 aspectos claves para tomar, de acuerdo con la literatura y la experiencia. Estos son el factor del ciclo ganadero, tomando como índice el precio y la dinámica del ganado hembra y cómo este afecta el macho, que en últimas es el que va principalmente a conformar la carne suministrada a los consumidores. El segundo factor tiene que ver con lo asociado directamente al precio del ganado macho entre ciudades, por medio de periodos anteriores cómo estos pueden afectar el precio en el largo plazo. Por último, la forma en que las precipitaciones como factor exógeno afectan el precio del ganado en ambas ciudades. (Lopez, 2019)

Es pertinente indicar que este proyecto de investigación busca a través de una nueva metaheurística disminuir el efecto negativo de los distintos factores que están limitando la efectividad y aplicabilidad de los distintos métodos de búsqueda en los problemas de optimización, específicamente en pronósticos de series de tiempo, tratando de alcanzar mejores respuestas y resultados con menor tiempo empleado.

## 1. MATERIALES Y MÉTODOS.

### 1.1. OPTIMIZACIÓN DE PRONÓSTICOS

#### 1.1.1. Ejercicios Problemas de Evaluación de Alta Complejidad

Para el desarrollo de la ruta esbozada, inicialmente se parte de la realización de ejercicios y análisis de Problemas de Evaluación de Alta Complejidad identificados en la literatura relacionada con pronósticos; con el objetivo de corroborar el potencial de trabajo del algoritmo cromático, con ocasión de la búsqueda de respuestas en problemas de optimización en series de tiempo.

#### 1.2. Caso de Estudio Practico (Precios del Ganado tipos de Córdoba)

En la segunda parte de los problemas propuestos para solucionar, se efectúa la estimación de precios para tipos de ganado seleccionados según datos históricos que se asocian con variables de tipo ganado referenciado como lo son; cantidad, mes y precios de venta. En esta oportunidad se conviene sobre hechos de venta obtenidos a través de información recolectada a partir de subasta ganadera Subastar S.A. de la ciudad de Montería- Córdoba.

El objeto aquí es implantar el algoritmo de búsqueda y el modelo no lineal de series de tiempo para realizar ejercicio de pronóstico en

tipos específicos de ganado, conforme a los datos escogidos.

### 1.3. Modelo de Regresión

Los modelos de regresión son usados para encontrar un mejor ajuste a un conjunto de datos. El propósito en todo tipo de regresión es tener la forma de la expresión (1). Para una serie de datos en donde  $Y$  es la variable dependiente medida para un experimento y  $X$  es la variable independiente que es cambiada durante los experimentos.

$$y = f(x) + \varepsilon_i \quad (1)$$

$F$  es una función que describe la relación de  $X$  y  $Y$ , y consta de uno o más parámetros.  $\varepsilon_i$  es la  $i$ -ésima observación del error que proviene de un nivel medio de datos que tiene una distribución normal con media de cero. El modelo que se utilizó para realizar el ajuste a los datos en los problemas que se trabajan en este artículo es el propuesto por (Behnamian & Ghomi, 2009) y se expone a continuación en las expresiones (2) y (3).

$$f(x) = \frac{f_1(x)}{c_1 + c_2 (f_1(x) + c_3 x \ln(b_0 + \sum_{i=1}^5 b_i x^{a_i}))} \quad (2)$$

En donde  $f_1(x)$  es:

$$f_1(x) = b_0 + \sum_{i=1}^5 b_i x^{a_i} + b_6^{a_6 x} + b_7 \sin(a_7 x) + b_8 \cos(a_8 x) \quad (3)$$

En donde los parámetros del modelo son:

$$b_0, b_1, b_2, b_3, b_4 \dots b_8, a_1, a_2, a_3, a_4 \dots a_8, c_1, c_2 \text{ y } c_3$$

El modelo establece la forma general de la ecuación de una regresión logística para prevenir los distintos casos de comportamientos en los datos. En este trabajo se realizó una adaptación del modelo anterior para poder trabajar con problemas de una y varias variables, el procedimiento es explicado a continuación.

En una serie de datos en donde se tiene que  $X_0, X_1, \dots, X_m$  son variables independientes que son cambiadas durante el experimento y  $Y$  es la variable dependiente medida para un



experimento se propone lo siguiente.

Para una variable  $X_i$  que va desde  $X_0$  hasta  $X_m$  se halla el valor  $f(x)$  correspondiente a las expresiones (2) y (3) y luego cada uno de los valores obtenidos para cada  $X_i$  se suman para obtener un  $\bar{F}_{st}$ , es decir que se tendría la expresión (4).

$$F_{sti} = f(X_0) + f(X_1) + \dots + f(X_m) \quad (4)$$

En donde el valor  $F_{sti}$  es hallado hasta el número  $n$  de observaciones. Luego entonces se calcula  $\bar{F}_{st}$  y  $\bar{Y}$  y con estos valores entonces se obtiene una constante de amplificación  $K$ . Este valor se obtiene mediante la expresión (5).

$$K = \frac{\bar{Y}}{\bar{F}_{st}} \quad (5)$$

Una vez encontrado  $K$  se obtiene el valor estimado  $\hat{Y}_i$  para cada variable dependiente  $Y_i$ , el valor estimado se calcula como se muestra a continuación en la expresión (6).

$$\hat{Y}_i = K F_{sti} \quad (6)$$

### 1.3.1. Optimización del Modelo

Generalmente los modelos no lineales son más complejos debido a que en estos poseen un gran número de parámetros, los cuales se deben hallar para encontrar el mejor ajuste a una serie de datos, es por eso que se trata de simplificar en lo posible la optimización de estos valores. Para obtener estos valores se suele recurrir a dos tipos métodos, el primero de estos es el desarrollo de métodos exactos para encontrar los valores óptimos de los parámetros del modelo y los segundos son procedimientos metaheurísticos para encontrar los valores de los parámetros que sean muy cercanos al óptimo del modelo (Da Silva, 2008). El segundo método es el más atractivo por que los procedimientos metaheurísticos no requieren de supuestos acerca de los parámetros (como es el caso de los métodos exactos) y ofrecen la posibilidad de

obtener un óptimo global de modelo expuesto en la expresiones (2) y (3), (Behnamian & Ghomi, 2009). Una vez conocidas la ventajas de los procedimientos metaheurísticos, se trabaja en base a estos y se utiliza un nuevo algoritmo de búsqueda llamado cromático desarrollado por (Sabie, 2011) el cual ofrece alternativas para encontrar muy buenos resultados en la búsqueda de los mejores parámetros para el modelo de regresión no lineal.

### 1.3.2. Función objetivo

Existen muchos índices de eficiencia para encontrar el mejor ajuste a una curva o una serie de datos, entre estos tenemos los indicadores de error como el ME, MAE, MSE y entre otros que son los más conocidos en medidas de eficiencia en pronósticos. Sin embargo, estos suelen tener algunas desventajas como la cancelación de errores de valores positivos con los de valores negativos como es el caso del ME (Tsolacos, 2006), catalogar a errores más grandes con un peso igual al de errores más pequeños como lo hace el MAE y el MSE (Rasmussen, 2004) y entre otros aspectos más, lo que nos llevó a un indicador de error más robusto que nos permitiera obtener resultados más precisos.

El indicador que se utilizó fue el  $R^2$  debido a que consultando en la literatura se encontró que este índice de error ofrecía los mejores resultados para estos tipos de problemas. Este índice se expone a continuación en la expresión (7).

$$R^2 = 1 - \frac{\sum_{i=0}^n (y_t - F_t)^2}{\sum_{i=0}^n (y_t - \bar{y}_t)^2} = \frac{SS_R}{SS_{total}} \quad (7)$$

En donde  $y_t$  es el valor observado y  $F_t$  es el valor ajustado,  $SS_{total}$  es la varianza de los datos y  $R^2$  es el valor real de la variabilidad que puede tener el modelo. Este índice expresa la proporción de varianza de la variable independiente explicada por la variable independiente (Eye & Schuster, 1998).

### 1.4. Codificación de Parámetros

Uno de los aspectos más importantes al trabajar con metaheurísticas, es decidir cómo se deben representar las soluciones y que tan eficientes

serian estas al realizar la búsqueda en el espacio de soluciones. Lo que se trata de hacer por lo general son codificaciones sencillas y que reduzcan el costo del algoritmo, debido a esto se decidió utilizar para la representación de todos los parámetros  $b_0, b_1, b_2, b_3, b_4 \dots b_8, a_1, a_2, a_3, a_4 \dots a_8, c_1, c_2$  y  $c_3$ . En la codificación de las soluciones del problema para el algoritmo cromático, la propuesta hecha por (Behnamian & Ghomi, 2009) cómo se muestra a continuación en la (ilustración N°2).

$b_0$	$b_1$	...	$b_8$	$a_1$	$a_2$	...	$a_8$	$c_1$	$c_2$	$c_3$
-------	-------	-----	-------	-------	-------	-----	-------	-------	-------	-------

**Ilustración 1** Representación de una solución en el algoritmo Cromático.

### 1.5. Algoritmo Cromático

Este es un nuevo método de optimización combinatoria que se inspira y hace una relación con las diferentes maneras con las que un músico u artista realiza y escoge la melodía más apropiada para una pieza musical, a partir de la combinación de distintas notas musicales de la escala cromática. La idea principal del método radica en que el músico elige o escoge la mejor melodía, la cual es la que mejor se adapta y representa apropiadamente una tonada específica.

Esta metaheurística se caracteriza por qué no necesita de un gran tamaño en el grupo de melodías que se tienen como opción de búsqueda, es decir, que no necesita de un gran número de soluciones iniciales para un obtener buenos resultados. El algoritmo trabaja con ML (melodía local) que es la mejor solución de la iteración y con MM (melodía global) la cual es mejor la solución encontrada hasta el momento, con estos realiza una búsqueda de vecinos a través de la escala cromática de las notas musicales y realiza movimientos de búsqueda que se fundamentan en conceptos musicales, además hace que la exploración del espacio de soluciones sea eficiente gracias a que utiliza la combinación de poblaciones con diferentes características, logrando así diversificar el conjunto de soluciones con el que se trabaja, es decir, que este algoritmo realiza arranques múltiples en su búsqueda para poder salir de óptimos locales, lo cual es una gran ventaja en

muchos problemas de optimización donde existen gran cantidad de variables.

No obstante, se resalta que esta metaheurística ofrece una muy buena alternativa para la recombinación y búsqueda de vecinos en problemas que son trabajados en codificación real (Sabie, 2011), lo que le da gran relevancia para obtener los mejores parámetros del modelo de las expresiones (2) y (3) debido a que estos tipos de problemas necesitan de este tipo de codificación. A continuación, se explican los principales operadores de búsqueda y los aspectos más importantes del algoritmo cromático.

#### 1.5.1. Vecinos de escala cromática

Para realizar la búsqueda de vecinos en números reales se adoptó la idea de la equivalencia de la recta numérica con la escala cromática, por lo que se le asignó un intervalo de números reales a cada nota musical en la recta numérica real, con una distancia de 0.5 entre un intervalo y otro, tomando como referencia el medio tono de distancia que guarda cada nota en la escala cromática musical. La equivalencia del número real con la nota musical de la escala se obtiene mediante las expresiones (8), (9) y (10).

$$\text{valor nota musical} = \beta - 12\alpha \quad (8)$$

En donde  $\beta$  es igual al redondeo del número real  $\delta$  a su entero siguiente y  $\alpha$  es igual a redondeo del número real  $\Omega$  a su entero anterior.

Donde tenemos que:

$$\delta = \frac{\text{numero real}}{0.5} \quad (9)$$

$$\Omega = \frac{\beta}{12} \quad (10)$$

De igual manera el grado para cada nota es igual al termino  $\alpha$ , dado el valor para el grado de cada nota se calcula el intervalo de búsqueda de vecinos para la nota SI mediante la expresión (11) y (12).

El límite inferior y superior del intervalo de la nota SI es:

$$L \text{ inf} = (\alpha)x_6 - 0.5$$

(11)

$$L_{sup} = L_{inf} + 0.5$$

(12)

Entonces el intervalo para la búsqueda de vecinos para la nota SI es igual  $((\alpha) \times 6 - 0.5, (\alpha) \times 6]$ .

Se aclara que este intervalo es abierto en el límite inferior y cerrado en el límite superior, no obstante, a partir de este intervalo de la nota SI se hallan los de las demás notas siguientes, sumando 0.5 a ambos límites del para cada nota. Entonces una vez identificado el valor de nota musical para cualquier número real, se obtiene su nota musical, luego con el grado específico de la nota se calculan los límites de nota inferiores y superiores para el intervalo de nota correspondiente.

Con los intervalos obtenidos para cada nota se podrían generar buenos resultados, pero debido a que estos son todavía muy grandes estos se dividen entre dos para realizar una búsqueda más minuciosa. Esta búsqueda y división es explicada a continuación. Si se desea realizar una búsqueda de vecinos en escala cromática del valor de la primera posición de la solución, se procede a encontrar la nota y el intervalo, estos se obtienen con las expresiones correspondientes (8), (9) y (10) junto con las expresiones (11) y (12).

De aquí se encuentra que el valor numérico que corresponde a la nota y su intervalo de búsqueda de vecino. Este intervalo se divide en dos y se obtiene dos nuevos intervalos como se expone a continuación.

El intervalo 1  $\rightarrow (L_{inf}, \frac{L_{sup}}{2}]$  y el intervalo 2

$$\rightarrow (\frac{L_{sup}}{2}, L_{sup}]$$

– Si el valor  $j$  de la posición de la mejor solución, a la que se le generara el vecino se encuentra en el intervalo 1.

– Entonces se generan dos números aleatorios así:

$s \rightarrow$  aleatorio entre  $L_{inf}$  y el valor  $j$   
(proveniente de una distribución uniforme)

$$p \rightarrow \text{aleatorio entre } j \text{ y el valor } \frac{L_{sup}}{2}$$

(proveniente de una distribución uniforme)

Una vez se obtienen los vecinos  $s$  y  $p$  se evalúa la solución con cada uno de ellos y se escoge la de mejor valor.

– Si en el caso contrario valor  $j$  de la posición de la mejor solución, a la que se le generara el vecino se encuentra en el intervalo 2.

Entonces se generan dos números aleatorios así:

$$s \rightarrow \text{aleatorio entre } \frac{L_{sup}}{2} \text{ y el valor } j$$

(proveniente de una distribución uniforme)

$$p \rightarrow \text{aleatorio entre } j \text{ y el valor del } L_{sup}$$

(proveniente de una distribución uniforme)

Una vez se obtienen los vecinos  $s$  y  $p$  se evalúa la solución con cada uno de ellos y se escoge la de mejor valor.

### 1.5.2. Vecinos de Inspiración y de Rotación

– VECINOS DE INSPIRACIÓN: Este tipo de variación resulta de un cambio pensamiento y sentido en la posición de una nota específica la cual es cambiada de forma imprevista y que busca mejorar la melodía que se tiene actualmente. Este tipo de vecinos se obtienen generando un número aleatorio  $j$  entre del número máximo de notas  $k$  de la melodía. El número  $j$  obtenido representa la nota y es la posición de la variable que se modificara a través de la inspiración, luego entonces lo que se hace es cambiar la nota escogida por un número aleatorio  $i$  del espacio de soluciones del problema y proveniente de una distribución uniforme.

– VECINOS DE ROTACIÓN DESCENDENTES: En los vecinos de rotación descendentes las variables de cada solución son consideradas notas a las cuales se les modificaran sus posiciones. Se obtiene la octava de la nota correspondiente. Esta se genera con un número aleatorio  $j$  entre 8 y las  $k$  variables del problema.

–  $8^{va} \rightarrow$  un número aleatorio  $j$  entre 8 y las  $k$  notas de la melodía conveniente actual



- Se obtiene la quinta de la nota correspondiente. Esta encuentra al restarle 3 posiciones a la 8<sup>va</sup>

5<sup>ta</sup> → se obtiene de → 8<sup>va</sup> - 3

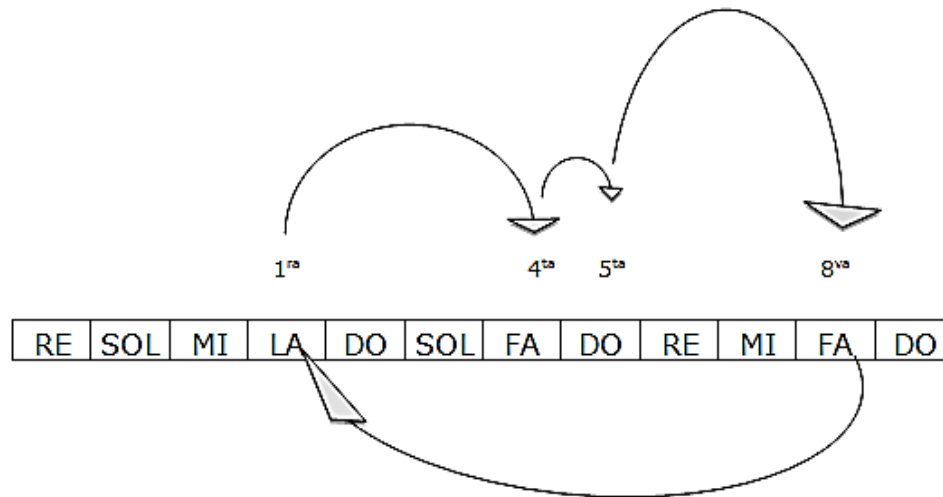
- Se obtiene la cuarta de la nota correspondiente. Esta encuentra al restarle 4 posiciones a la 8<sup>va</sup>

4<sup>ta</sup> → se obtiene de → 8<sup>va</sup> - 4

- Se obtiene la primera de la nota correspondiente. Esta encuentra al restarle 7 posiciones a la 8<sup>va</sup>

1<sup>ra</sup> → se obtiene de → 8<sup>va</sup> - 7

Los movimientos para la generación de vecinos de rotación descendentes son expuestos a continuación en la (ilustración N°3). Cada nota gira hacia delante tomando la posición clave más cercana.



**Ilustración 2** Vecinos de Rotación descendentes

#### – VECINOS DE ROTACIÓN ASCENDENTES:

En los vecinos de rotación ascendentes las variables de cada solución son consideradas notas a las cuales se les modificaran sus posiciones. Se obtiene la primera de la nota correspondiente. Esta se genera con un número aleatorio  $j$  entre 1 y las  $(k-8)$  variables del problema.

1<sup>ra</sup> → un número aleatorio  $j$  entre 1 y las  $k-8$  posiciones de la melodía conveniente actual

- Se obtiene la cuarta de la nota correspondiente. Esta encuentra al sumarle 3 posiciones a la 1<sup>ra</sup>

4<sup>ta</sup> → se obtiene de → 1<sup>ra</sup> + 3

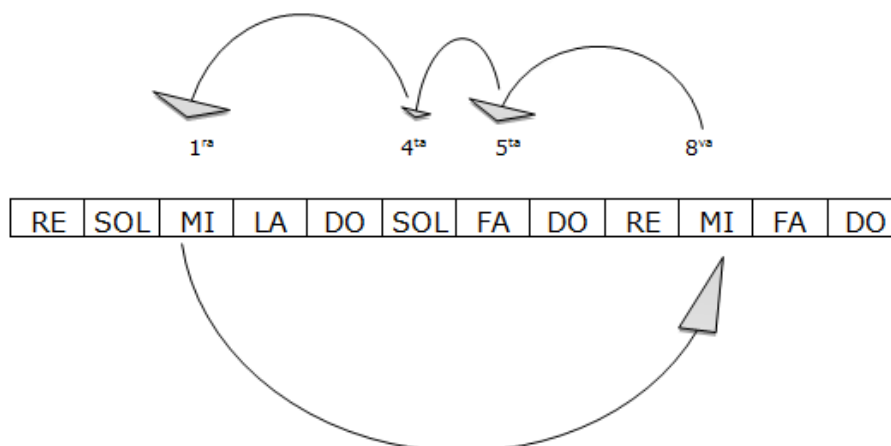
- Se obtiene la quinta de la nota correspondiente. Esta encuentra al sumarle 4 posiciones a la 1<sup>ra</sup>

5<sup>ta</sup> → se obtiene de → 1<sup>ra</sup> + 4

- Se obtiene la octava de la nota correspondiente. Esta encuentra al sumarle 7 posiciones a la 1<sup>ra</sup>

8<sup>va</sup> → se obtiene de → 1<sup>ra</sup> + 7

Los movimientos para la generación de vecinos de rotación ascendentes son expuestos a continuación en la (ilustración N°4). Cada nota gira hacia atrás tomando la posición clave más cercana.



**Ilustración 3** Vecinos de Rotación ascendentes.

## 1.6. Parámetros y Pseudocódigo del Algoritmo Cromático

### 1.6.1. Parámetros Evaluación De Ejercicios Alta Complejidad

Para esta primera parte se trabajó con un tamaño 10 en el grupo de melodías iniciales por ARRAN, este parámetro ARRAN fue 100 iteraciones y el criterio máximo de parada fue de 1000 iteraciones. El valor para AM fue de 0.0971352, MC, VMC y UM fueron de 0.552618 0.233438 y 0.483767 respectivamente (Sabie, 2011). Para todos los parámetros:

$b_0, b_1, b_2, b_3, b_4 \dots b_8, a_1, a_2, a_3, a_4 \dots a_8, c_1, c_2$  y  $c_3$  Se realizaron sus estimaciones en un rango límite de  $[-1,1]$ . Para los parámetros  $b_0, c_1, c_2$  se realizaron estimaciones iniciales de 1, 1 y 0 respectivamente para prevenir indeterminaciones dentro del modelo.

### 1.6.2. Parámetros Evaluación Caso de Estudio Practico

En esta oportunidad se realizaron corridas con un tamaño 10 en el grupo de melodías iniciales por ARRAN, este parámetro ARRAN fue 30 iteraciones y el criterio máximo de parada fue de 8 horas. El valor para AM fue de 0.0971352, MC, VMC y UM fueron de 0.552618 0.233438 y 0.483767 respectivamente (Sabie, 2011). Para todos los parámetros:

$b_0, b_1, b_2, b_3, b_4 \dots b_8, a_1, a_2, a_3, a_4 \dots a_8, c_1, c_2$  y  $c_3$  Se realizaron sus estimaciones en un rango

límite de  $[-1,1]$ . Igualmente, para los parámetros  $b_0, c_1, c_2$  se realizaron estimaciones iniciales de 1, 1 y 0 respectivamente para prevenir indeterminaciones dentro del modelo.

## 2. RESULTADOS Y DISCUSIONES

En esta sección inicialmente se compara el algoritmo propuesto con varios modelos de tipo lineal y no lineal, en unos cuatro problemas que pertenecen al grupo de modelos “test problems” los cuales son considerados por el National Institute Standards and Technology USA website, como modelos con un alto nivel de dificultad.

Posteriormente son mostrados los resultados del algoritmo en unos problemas multivariados en caso de estudio práctico. Estos mismos se desprenden de la valoración y análisis de la estimación de precios para tipos de ganado seleccionados según datos históricos que se asocian con variables de tipo ganado, referenciado como lo son; cantidad, mes y precios de venta. Cabe resaltar que estos se obtienen sobre hechos de venta e históricos. Los mismos datos fueron obtenidos a través de información generada por la empresa Subastar S.A. de la ciudad de Montería- Córdoba.

En específico se implementó el algoritmo de búsqueda y el modelo no lineal de series de tiempo para realizar los ejercicios de pronóstico en tipos específicos de ganado, conforme a los datos recolectados. A continuación, se presenta

el pseudocódigo e información referente a cerca del diseño del algoritmo de búsqueda.

## Pseudocódigo del algoritmo.

```

INICIALIZAR LOS PARÁMETROS INCLUYENDO, TAMAÑO(N) DE GRUPO DE MELODÍAS INICIALES(GMI)
POR ARRAM ,MÁXIMO DE ITERACIONES POR ARRAM, AM, UM, MC Y VMC
MIENTRAS QUE NO SE ALCANCE LA CONDICIÓN DE PARADA
    MIENTRAS QUE NO SE ALCANCE EL MÁXIMO DE ITERACIONES POR ARRAM
        GENERAR GRUPO DE MELODÍAS INICIALES (GMI) DE TAMAÑO(N)
        PARA i = 1 HASTA EL TAMAÑO (N) DEL GRUPO DE MELODÍAS(GMI)
            EVALUAR CADA  $X_i$  MELODÍA DEL GRUPO
            SI FITNESS DE ( $X_i$ ) ES MEJOR QUE EL FITNESS (ML)
                 $ML \leftarrow X_i$  ; FITNESS (ML)  $\leftarrow$  FITNESS ( $X_i$ )
            FIN SI
            SI FITNESS DE (ML) ES MEJOR QUE EL FITNESS (MM)
                 $MM \leftarrow ML$ ; FITNESS (MM)  $\leftarrow$  FITNESS (ML)
            FIN SI
        FIN PARA
    PARA i = 1 HASTA EL TAMAÑO (N) DEL GRUPO DE MELODÍAS
        SI LA PROBABILIDAD AM ES MAYOR QUE UN NUMERO ALEATORIO
            ENTRE CERO Y UNO PARA LA  $X_i$  MELODÍA
                PARA j = 1 HASTA EL NUMERO K DE VARIABLES DE LA MELODÍA  $X_i$ 
                    SI LA PROBABILIDAD UM ES MAYOR QUE UN NUMERO ALEATORIO
                        ENTRE CERO Y UNO
                            REPLAZAR EL VALOR DE LA VARIABLE j DE LA melodía  $X_i$ 
                            POR EL VALOR DE LA VARIABLE j DE LA MM
                    SI NO
                        REALIZAR UNA BÚSQUEDA DE VECINOS DE ESCALA CROMÁTICA
                        PARA EL VALOR DE LA VARIABLE j DE LA MELODÍA  $X_i$ 
                    FIN SI
                FIN PARA
            SI NO
                SI LA PROBABILIDAD MC ES MAYOR QUE UN NUMERO ALEATORIO
                    ENTRE CERO Y UNO PARA LA  $X_i$  MELODÍA
                        SE REPLAZA LA  $X_i$  MELODÍA POR LA ML
                SI NO
                    ESCOGER DOS MELODÍAS ALEATORIAS (S, P) DEL GRUPO DE MELODÍAS
                    ASIGNAR EL VALOR UNO A LA MELODÍA (S) Y DOS A LA MELODÍA (P)
                    PARA j = 1 HASTA EL NUMERO K DE VARIABLES DE LA MELODÍA  $X_i$ 
                        GENERAR UN NÚMERO ALEATORIO ENTRE UNO Y DOS
                        SI NUMERO ALEATORIO ES IGUAL A UNO
                            REPLAZAR EL VALOR DE LA VARIABLE j DE LA MELODÍA  $X_i$ 
                            POR EL VALOR DE LA VARIABLE j DE LA MELODÍA S
                        SI NO
                            REPLAZAR EL VALOR DE LA VARIABLE j DE LA MELODÍA  $X_i$ 
                            POR EL VALOR DE LA VARIABLE j DE LA MELODÍA P

```

```

                FIN SI
            FIN PARA
        FIN SI
    SI    LA PROBABILIDAD VMC ES MAYOR QUE UN NUMERO ALEATORIO
        ENTRE CERO Y UNO PARA LA  $X_i$  MELODÍA
        SE REALIZA UNA BÚSQUEDA DE VENOS DE INSPIRACIÓN
    SI NO
        SE REALIZA UNA BÚSQUEDA DE VENOS DE ROTACIÓN DE NOTAS
    FIN SI
FIN SI

    FIN PARA
FIN PARA
FIN PARA
FIN
```

### 2.1.1. Problemas de Evaluación: Ejercicios de Alta Complejidad

#### 2.1.2. Problema de evaluación N°1 (Clase polinomial)

Este problema de evaluación tiene 82 datos y es conocido como “Fillip” y consta de 11 parámetros, el modelo es expuesto en la expresión (13).

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_9 x^9 + \beta_{10} x^{10} + \varepsilon \quad (13).$$

Los resultados del modelo propuesto son expuestos en el (Gráfico N°1) y en la (Tabla N°1) se presenta la comparación de resultados de la solución estándar y del modelo propuesto.

#### 2.1.3. Problema de evaluación N°2 Clase (polinomial)

Este problema de evaluación tiene 21 datos, es conocido como “Wampler4” y consta de 6 parámetros, el modelo es expuesto en la expresión (14).

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4 + \beta_5 x^5 + \varepsilon \quad (14).$$

Los resultados del modelo propuesto son expuestos en el (Gráfico N°2) y en la (Tabla N°2) se presenta la comparación de resultados de la solución estándar y del

modelo propuesto.

#### 2.1.4. Problema de evaluación N°3 (Clase racional)

Este problema de evaluación tiene 11 datos, es conocido como “MGH09”, y consta de 4 parámetros, el modelo es expuesto en la expresión (15).

$$y = \frac{\beta_1 (x^2 + x\beta_2)}{x^2 + x\beta_3 + \beta_4} \quad (15).$$

Los resultados del modelo propuesto son expuestos en el (Gráfico N°3) y en la (Tabla N°3) se presenta la comparación de resultados de la solución estándar y del modelo propuesto.

#### 2.1.5. Problema de evaluación N°4 (Clase diversa)

Este problema de evaluación tiene 154 datos, es conocido como “Bennett5”, y consta de 4 parámetros, el modelo es expuesto en la expresión (16).

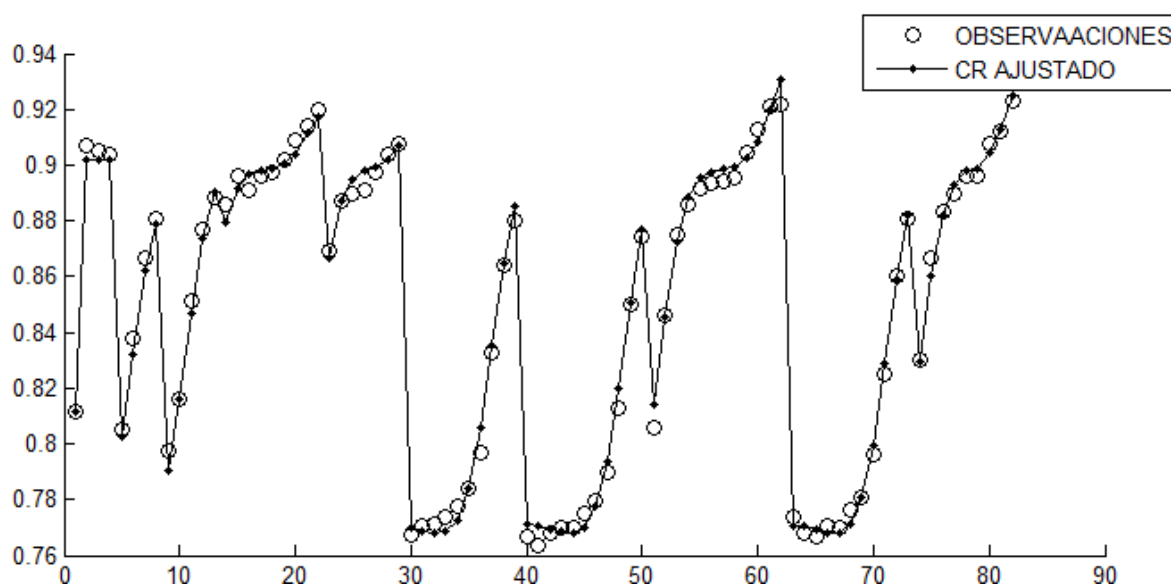
$$y = \beta_1 (x + \beta_2)^{-1/\beta_3} + \varepsilon \quad (16).$$

Los resultados del modelo propuesto son expuestos en el (Gráfico N°4) y en la (Tabla N°4) se presenta la comparación de resultados de la solución estándar y del modelo propuesto.



**Tabla 1** Comparación de resultados problema N°1.

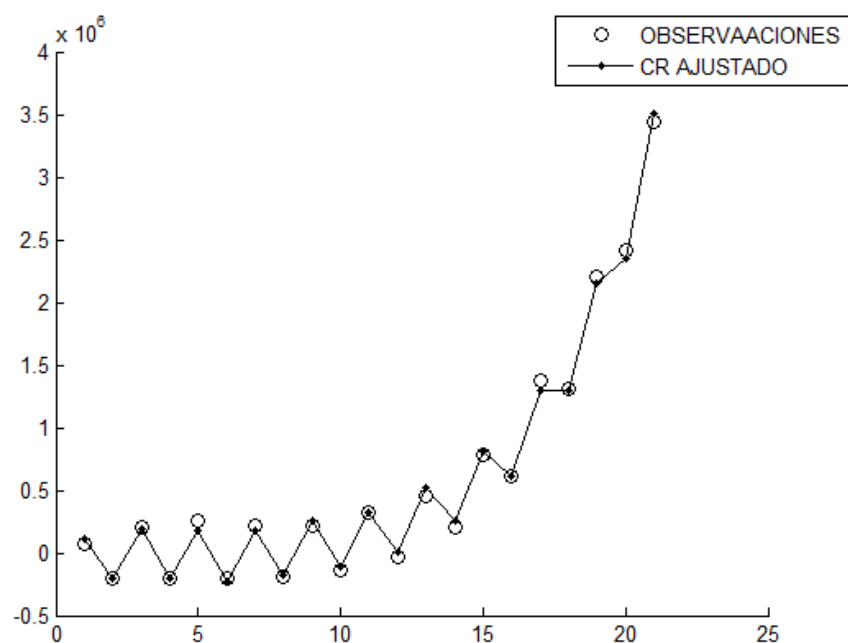
MÉTODO	R- CUADRADO	PREDICCIÓN DE ERROR
SOLUCIÓN ESTÁNDAR	0.996	0.004
MODELO PROPUESTO	0.995	0.005



**Gráfico 1** Ajuste para el problema N°1

**Tabla 2** Comparación de resultados problema N°2.

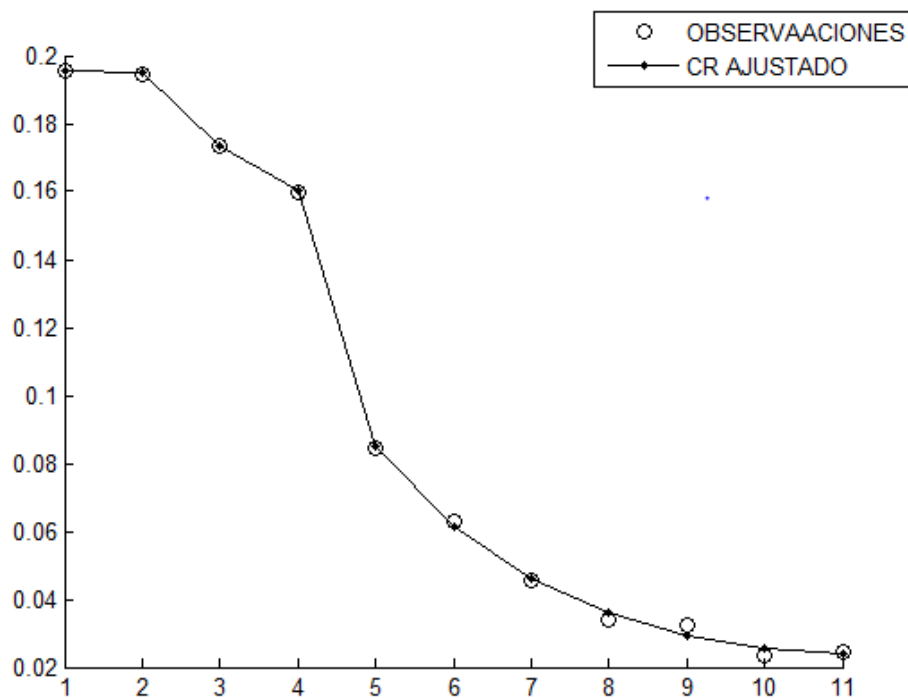
MÉTODO	R- CUADRADO	PREDICCIÓN DE ERROR
SOLUCIÓN ESTÁNDAR	0.9574	0.0426
MODELO PROPUESTO	0.9982	0.0018



**Gráfico 2** Ajuste para el problema N°2

**Tabla 3** Comparación de resultados problema N°3.

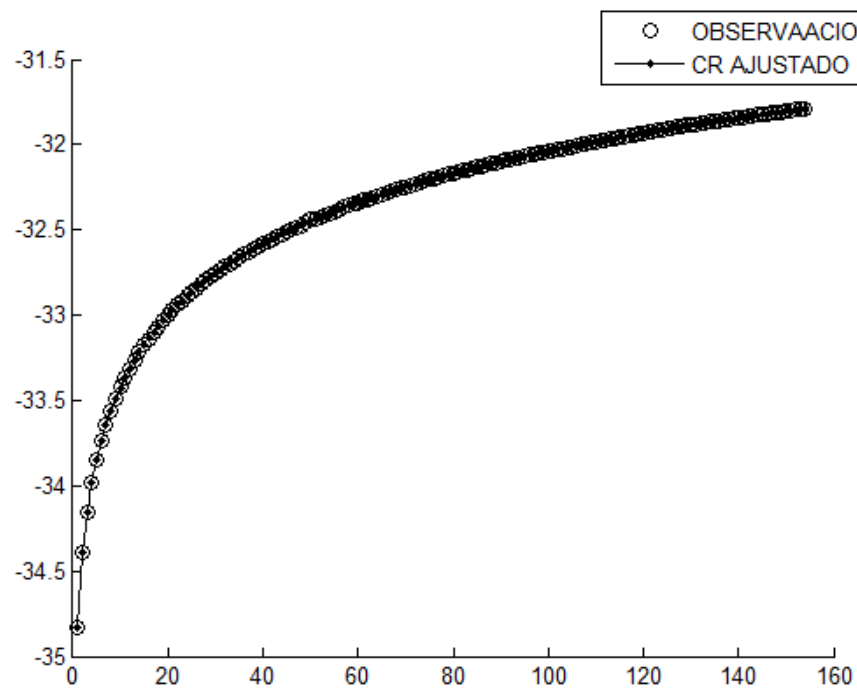
MÉTODO	R- CUADRADO	PREDICCIÓN DE ERROR
SOLUCIÓN ESTÁNDAR	0.9915	0.0025
MODELO PROPUESTO	0.9996	0.0004



**Gráfico 2** Ajuste para el problema N°3.

**Tabla 4** Comparación de resultados problema N°4.

MÉTODO	R- CUADRADO	PREDICCIÓN DE ERROR
SOLUCIÓN ESTÁNDAR	0.99998	0.00001
MODELO PROPUESTO	0.99998	0.00001



**Gráfico 3** Ajuste para el problema N°4.

## **2.2. Problemas con Múltiples Variables**

### **3.2.1. Caso Precios del Ganado Tipo Hembra de Vientre 2 ¾ (HV) en Córdoba**

Este ejercicio de evaluación de se procede a realizar la estimación de precios para ganado tipo hembra de vientre, edad de 2 ¾ años, HV, a partir de las variables tiempo, cantidad, de la siguiente manera; mes del año, cantidad promedio mensual, para estimar el valor del precio promedio de venta del mes por kilogramo.

Importante señalar que la estimación se lleva a cabo con fundamento en los datos históricos recopilados para los años 2014-2019. La fuente directa de información se concreta gracias al contenido público disponible en la web oficial la empresa Subastar S.A. de la ciudad de Montería- Córdoba.

Los datos y resultados del modelo propuesto para la estimación de precios ganado tipo HV son expuestos en las (Tabla N°6) y en el (Gráfico N°5).

### **3.2.2. Caso Precios del Ganado Tipo Vaca Escotera 3 (VE) en Córdoba**

En este punto se estiman los precios para ganado tipo vaca escotera con 3 años de edad, VE, a partir de las variables tiempo, cantidad, de la siguiente manera; mes del año, cantidad promedio mensual, para estimar el valor del precio promedio de venta del mes por kilogramo.

La estimación se lleva a cabo con fundamento en los datos históricos recopilados para los años 2014-2019. La fuente directa de información se concreta gracias al contenido público disponible en la web oficial la empresa Subastar S.A. de la ciudad de Montería- Córdoba.

Los datos y resultados del modelo propuesto para la estimación de precios ganado tipo VE son en las (Tabla N°7) y en el (Gráfico N°6).

### **3.2.3. Caso Precios del Ganado Tipo Toro 3 (TO) en Córdoba**

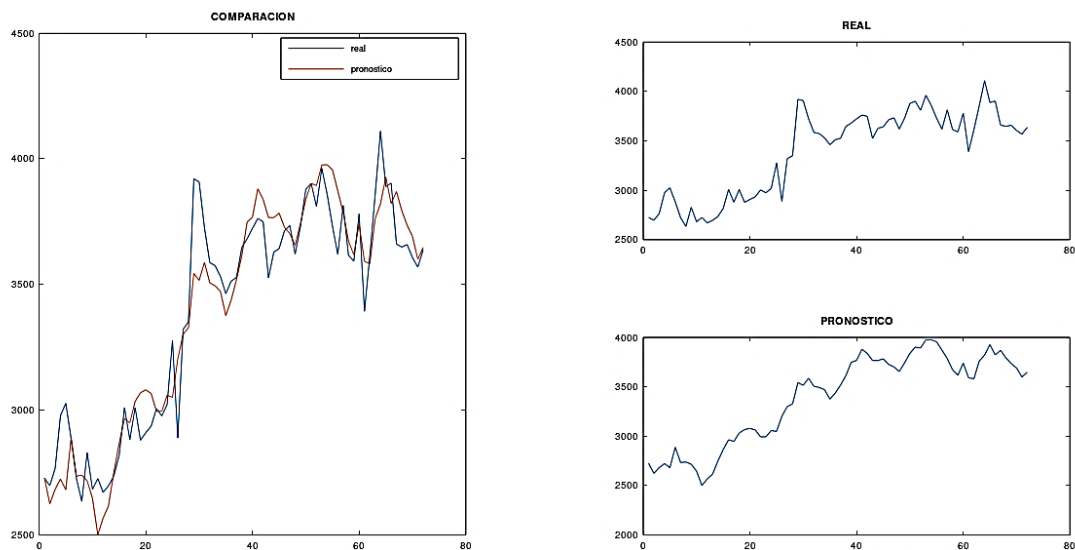
En el caso de esta aplicación se estiman los precios para ganado tipo toro de 3 años, TO, a partir de las variables tiempo, cantidad, de la siguiente manera; mes del año, cantidad promedio mensual, para estimar el valor del precio promedio de venta del mes por kilogramo.

La estimación se lleva a cabo con fundamento en los datos históricos recopilados para los años 2014-2019. La fuente directa de información se concreta gracias al contenido público disponible en la web oficial la empresa Subastar S.A. de la ciudad de Montería- Córdoba.

Los datos y resultados del modelo propuesto para la estimación de precios ganado tipo TO son en las (Tabla N°8) y en el (Gráfico N°7).

**Tabla 5** Comparación de resultados ganado tipo HV.

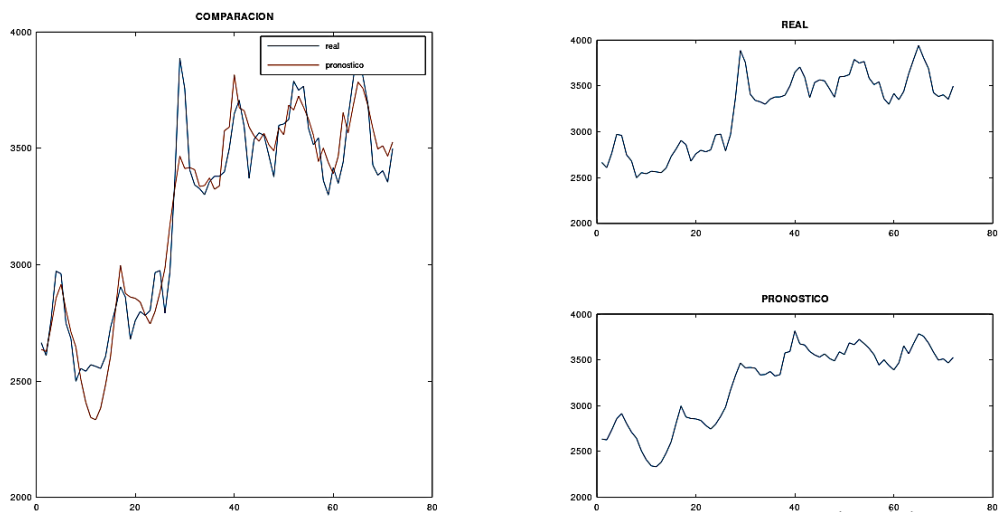
MÉTODO	R- CUADRADO	PREDICCIÓN DE ERROR
MODELO PROPUESTO	0.913229116930542	0.0867708830694582



**Gráfico 4** Ajuste para el caso tipo HV

**Tabla 7** Comparación de resultados ganado tipo VE.

MÉTODO	R- CUADRADO	PREDICCIÓN DE ERROR
MODELO PROPUESTO	0.92336262	0.07663738

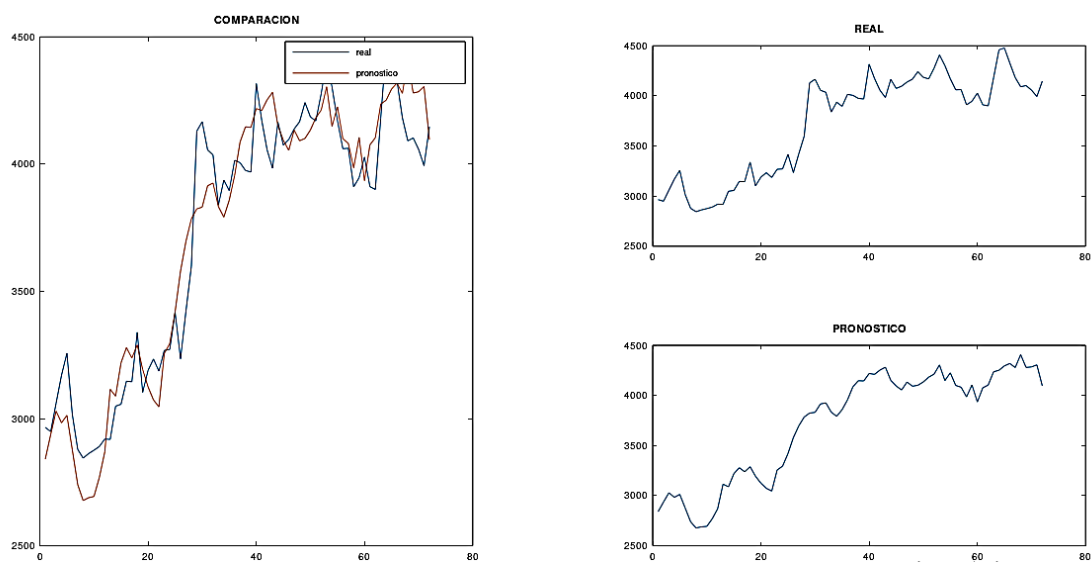




**Gráfico 5** Ajuste para el caso tipo VE.

**Tabla 8** Comparación de resultados ganado tipo TO.

MÉTODO	R- CUADRADO	PREDICCIÓN DE ERROR
MODELO PROPUESTO	0.92690734	0.07309266



**Gráfico 6** Ajuste para el caso tipo TO.

#### 4. CONCLUSIONES

En este proyecto se trabajó con un modelo de regresión no lineal en pronóstico de series de tiempo con diferentes problemas de gran dificultad. Se encontró que el algoritmo cromático ofrece muy buenos resultados en la búsqueda, de conformidad con los parámetros para el modelo.

De los ejercicios desarrollados, se destaca a este algoritmo debido a que ofrece grandes ventajas en la exploración y la diversificación de las soluciones, lo que es muy importante a la hora de encontrar óptimos globales en un espacio de soluciones. Los resultados obtenidos para los problemas de alta dificultad muestran que se supera en la mayoría de los casos las soluciones estándares encontradas en la literatura.

Para los ejercicios relacionados con los casos prácticos se identificó que el modelo arroja buenos resultados para la estimación de valores a partir de múltiples variables. En esta propuesta se evidencia la eficiencia y robustez del modelo propuesto para encontrar muy buenos ajustes en diferentes series de datos, tanto de una, como de múltiples variables. El modelo se adecuó satisfactoriamente y tuvo un alto poder predictivo en comportamientos lineales, no lineales, cíclicos y estacionales.

Además, hay que resaltar que en el algoritmo también logra que este sea capaz de hacer que sus soluciones cumplan con los supuestos o restricciones que son por lo general necesarias para hacer uso de los modelos de regresión, lo que ratifica a un mas esta nueva propuesta y la hace más sólida y confiable.

#### REFERENCIAS

- [1]. Barreras Serrano, A., Eduardo, S. L., & Figueroa Saavedra, F. (2014). Uso de un modelo univariado de series de tiempo para la predicción, en el corto plazo, del comportamiento de la producción de carne de bovino en Baja California, México. \*Instituto de Investigaciones en Ciencias Veterinarias de la Universidad Autónoma de Baja California.
- [2]. Behnamian, J., & Ghomi, F. (2009). Development of a PSO-SA hybrid metaheuristic for a new comprehensive. Elsevier.
- [3]. Botero, S., & Cano, J. (2008). Análisis de series de tiempo para la predicción de los precios de la energía en la bolsa de Colombia. CUADERNOS DE ECONOMÍA 48, 2008 .
- [4]. Castillo, O. (2007). Comportamiento temporal de los precios del ganado macho de levante de primera en Sincelejo. Temas Agrarios Universidad de Córdoba, Departamento de Ingeniería Agronómica y Desarrollo Rural. .
- [5]. Castillo, O. (2007). Estacionalidad, ciclos y volatilidad en los precios del ganado macho de levante en Montería, Colombia. Rev MVZ, Cordoba.
- [6]. Espinosa, J., Vélez, & A. (2017). Descripción y pronóstico de la producción de carne de bovino en el estado de Tabasco. Actas Iberoamericanas en Conservación Animal .
- [7]. Eye, A. V., & Schuster, C. (1998). Regression analysis for social sciences (1st ed.). . San Diego: Academic Press.
- [8]. Lopez, S. (2019). Determinantes del precio del ganado gordo bovino en pie de Medellín y Bogotá: 2009-2019. Medellín, Colombia: Escuela De Economía Y Finanzas - Universidad EAFIT.
- [9]. Montgomery, D. C. (2004). Desing

- and analysis of experiments book. 458-459.
- [10]. Rasmussen, R. (2004). On time-series data and optimal parameters. . Omega, 32,111–120.
- [11]. Sabie, R. (2011). Un nuevo método de optimización que se fundamenta a través de un algoritmo de búsqueda basado en la escala cromática de las notas musicales. Montería, Cordoba, Colombia: Universidad de Córdoba.
- [12]. Tsolacos, S. (2006). An assessment of property performance forecasts, consensus versus econometric. . Journal of Property Investment & Finance, 24,386–399.